

Kyligence Analytics Platform

2017产品白皮书





世界领先的开源 OLAP 分析引擎

Apache Kylin 是一个开源的分布式分析引擎, 提供 Hadoop 之上超大规模数据的 SQL 查询接口及多维分析能力。在超大规模数据集上建立数据模型, 构建支持多维分析的预计算 Cube, 并开放通用的 ODBC、JDBC 或 Restful API 接口。这种独特的预计算 Cube 使 Apache kylin 可以轻松应对超大数据集上的查询, 并实现亚秒级响应。

Kylin: 首个由中国团队贡献的 Apache 顶级项目

Kylin 源起于 eBay 中国, 于 2015 年 11 月正式毕业成为 Apache 软件基金会 (ASF) 顶级项目, 是第一个由中国团队完整贡献到 Apache 软件基金会的顶级项目。

Apache Kylin全球用户



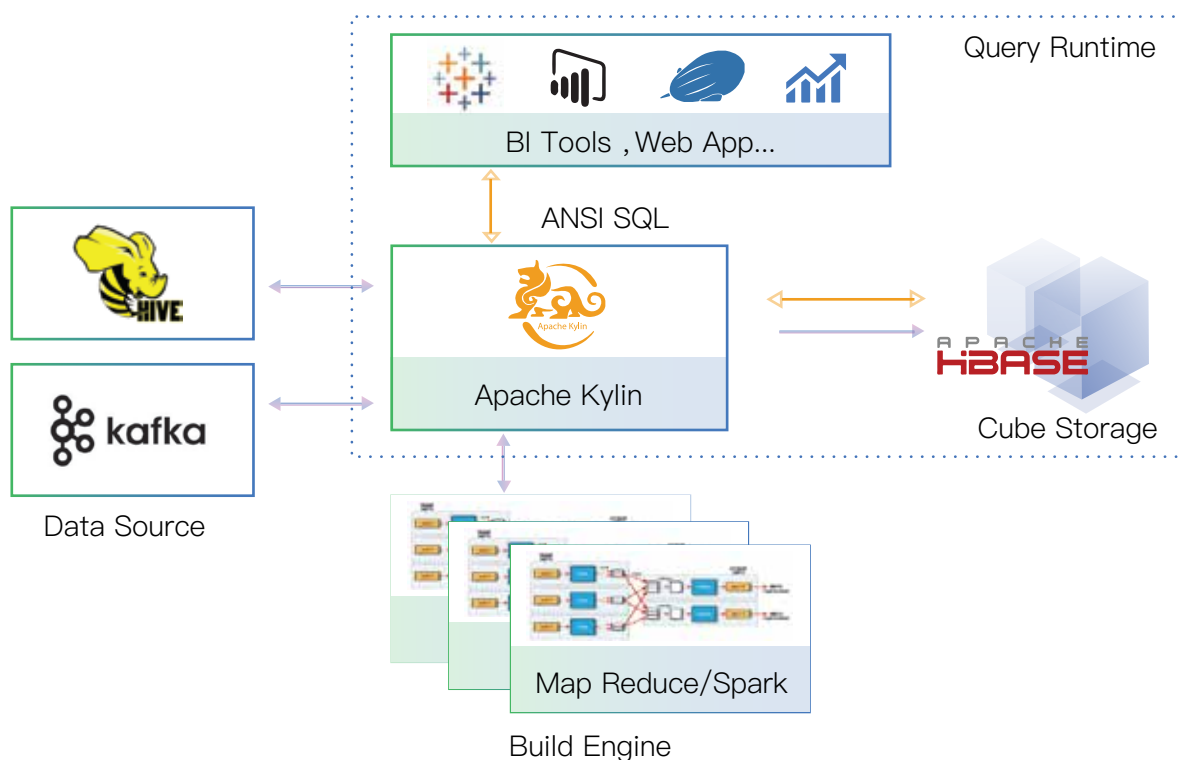
Apache Kylin 满足了 eBay 多种经典数据产品需求: 用户行为分析报告, 流量分析报告, 搜索引擎优化等。 --- eBay

满足了海量数据场景下交互式分析的挑战, 支撑了万亿规模数据集的秒级交互式分析及数千分析师的日程分析需求。 --- 今日头条

支持 4G 流量业务分析, 满足了每天处理两千亿条记录的业务需求。 --- 中国移动

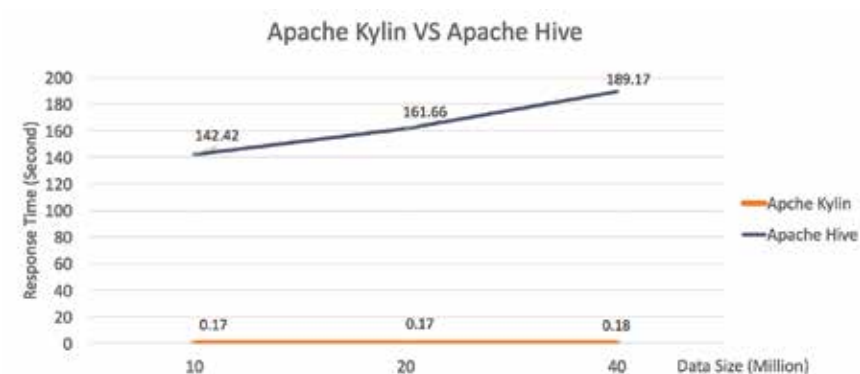
有效降低了 hadoop 上数据仓库的构建时间和成本, 自助分析平台大大提升了分析师的生产效率。 --- OPPO

系统架构



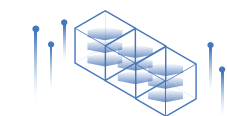
性能对比

- 对比 Hive, 查询加速百倍以上
- 常量级查询延迟, 应对大数据增长, 提供亚秒级响应
- 超高 QPS, 支持并发能力线性扩展



* 基于 Star Schema Benchmark 数据集进行对比测试

核心优势



亚秒级高性能查询



标准 SQL



非侵入式架构



支持流式构建



Kyligence Analytics Platform

Kyligence Analytics Platform (KAP) 是 Kyligence 提供的基于 Apache Kylin 的企业级大数据智能分析平台，在 PB 级数据集上提供亚秒级标准 SQL 查询响应，提供互联网级的高并发访问，赋能分析师以行业标准的数据仓库和商业智能方法论架构基于 Hadoop 的解决方案。KAP 为商业用户、分析师和工程师提供了统一的分析平台，支持自助式建模，无需编程，并与主流 BI 工具实现无缝集成，在开源 Apache Kylin 核心功能之外，在企业用户所关注的实施效率、安全控制、性能优化、自助式敏捷 BI、系统监控和管理等方面进行了全面创新和增强。作为 Hadoop 上的原生 OLAP 解决方案，KAP 基于 Hadoop 标准接口与集群交互，兼容主流 Hadoop 发行版，支持私有数据中心及云端部署安装。

BI 可视化

交互式分析

报表

仪表盘

数据挖掘

数据探索

OLAP

数据集市

Kyligence Analytics Platform

– powered by Apache Kylin

大数据平台

HDFS

YARN

Map
Reduce

Hive

Spark

Kafka

...

按需部署

本地部署

混合部署

云端部署

KAP: Apache Kylin企业版



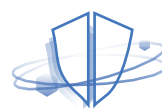
统一大数据
查询入口



智能化的
可视建模



无缝集成 BI
零学习门槛



增强的
企业级特性



统一大数据查询入口

采用Hybrid OLAP融合式架构，超越OLAP，同时满足性能关键的亚秒级的查询响应，也支持灵活的探索式分析，为用户提供统一的大数据查询入口。

查询下压

通过智能查询路由，将 Cube 不支持的查询下压到其他 SQL 引擎，内置支持 Spark SQL 及 Hive，可以同时满足对性能有严格要求的报表类应用，以及灵活的探索式分析。

多路索引算法

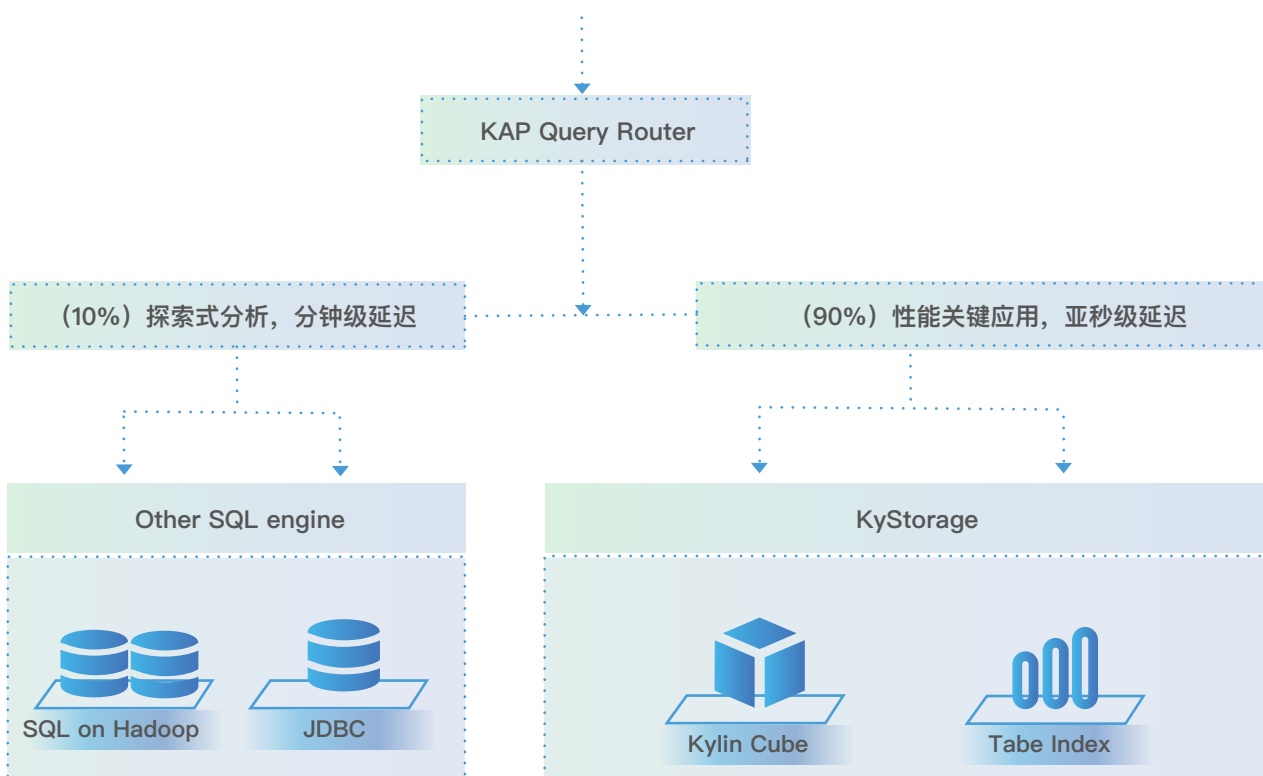
针对分析优化的索引结构，可进行多路复合索引，倒排索引，针对超高基数维度、复杂过滤条件、明细查询等场景进行了专门的优化。

快速列式存储引擎

基于 HDFS 的列式存储引擎 KyStorage，针对分析场景优化数据访问效率，相对 Apache Kylin，查询性能有几倍到几十倍的提升，在存储空间上节省超过 50%。

明细查询

全面支持明细数据的查询，优化了对宽表的支持，降低了数据建模的难度，更好地服务数据探索式分析场景。





智能建模平台

融入多项智能辅助能力，为模型设计人员、分析师提供一站式数据建模建议，提高生产力，赋能分析师。

智能构建

智能推荐维度、度量组合，模型一键优化，Cube 智能剪枝，帮助分析师降低学习成本，令 Cube 更智能，再降低 Cube 的膨胀率的同时，提升查询和构建的效率。

建模中心

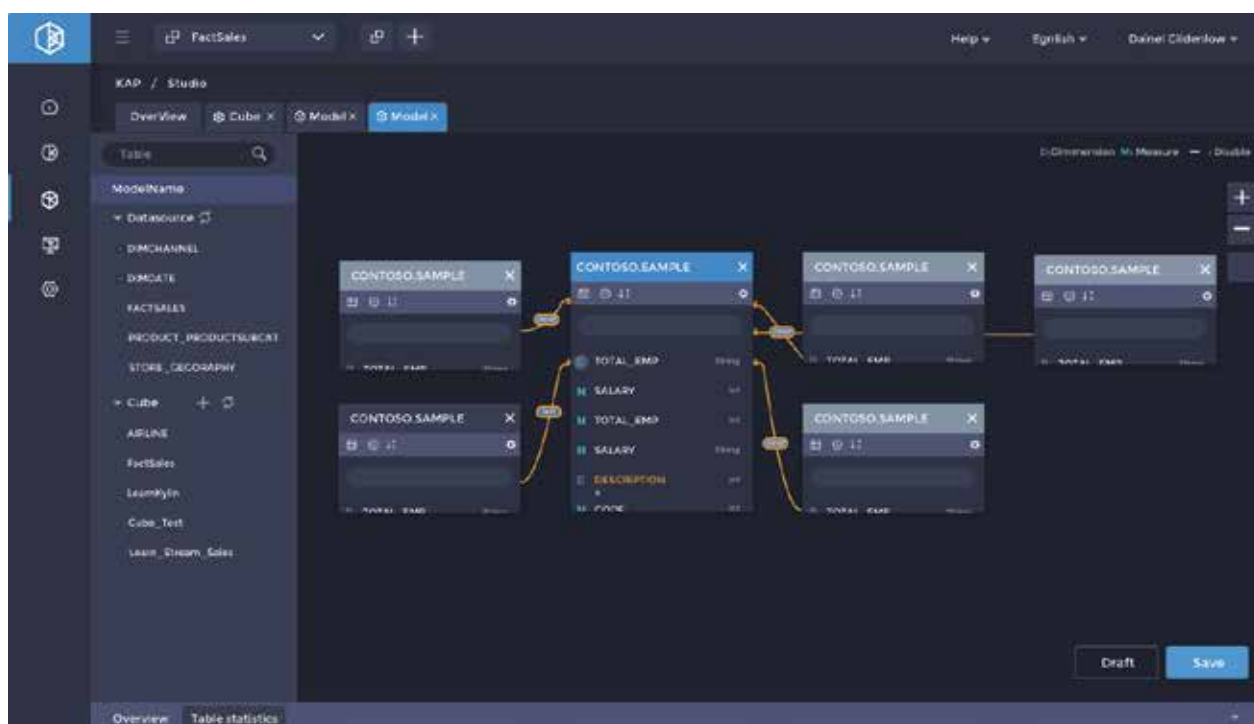
可视化建模组件 KyStudio，拖拽式数据建模流程，直观导入数据源；支持行业标准的星型模型，雪花模型；模型健康度检测帮助用户及时发现模型设计缺陷。

预置多种分析函数

在传统统计函数之外，预置 Top-N、Distinct Count、Percentile 等多种数据分析常用函数，针对超大数据规模优化算法，助力分析师提高分析效率和能力。

增强的语义层

支持将 Hive 内置函数及 UDF 等操作定义为可计算列，有效实现模型降维，数据清理。增强模型语义层表达，助力分析师自助式分析，减少 ETL 依赖。





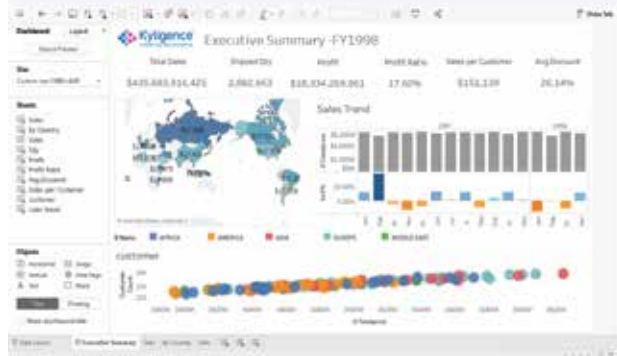
无缝整合 BI

BI 整合浑然一体, 令分析真正做到所见即所得

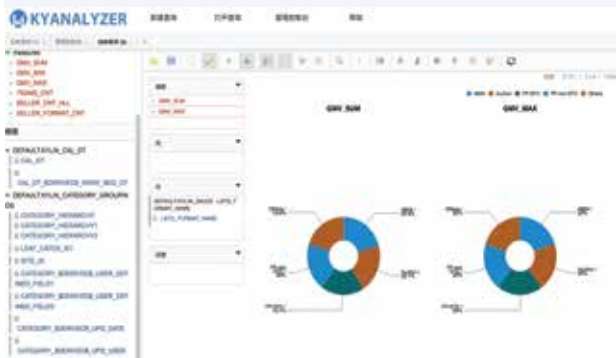
• Power BI



• Tableau



• KyAnalyzer



• 更多



排名不分先后

无缝集成主流 BI

与第三方 BI 工具无缝集成, 支持 Tableau、Cognos、MicroStrategy、Excel、PowerBI 等主流 BI 厂商, 令业务部门实现数据加速的同时, 也可沿用部门原有 BI 投资。

内置开箱即用的敏捷BI工具

内置敏捷 BI 工具 KyAnalyzer, 实现一站式分析, 深度集成 KAP 数据模型与权限控制, 提供多样分析解决方案。

通用 BI 接口

提供 ODBC / JDBC 等通用的数据驱动接口, 轻松对接不同的可视化及数据分析工具, 使用业务人员, 分析师熟悉的 SQL 语言或 BI 工具进行大数据分析。

分析师零学习即可上手

所见即所得的操作界面, 行业通用的数据建模方法论等特性, 令分析师可以轻松上手, 零学习门槛。



更多企业级特性

为企业客户设计，提升可靠性、可管理能力，DevOps友好，与IT系统完美整合。

兼容性

兼容开源，与 Cloudera、Hortonworks、华为等主流 Hadoop 发行商实现产品双向认证，令首次部署完美融合当前的 IT 环境。可支持数据中心及云端部署。

细粒度权限控制

支持项目级，表级，单元格级等细粒度权限控制，支持针对不同部门，不同组织的用户的最细粒度访问控制。

增强的安全管理

支持 LDAP、SSO，让 KAP 轻松和您的 IT 系统进行完美整合，助力 IT 部门进行统一的安全管控。

元数据管理

简化的安装升级流程，增强的元数据存储方式，多种备份方式，可使用关系型数据库备份元数据，令系统运维更轻松。

	Apache Kylin	KAP	KAP Plus
定位	基于 Hadoop 的多维分析平台	企业级 Hadoop 多维分析平台	企业级 Hadoop 多维分析平台
系统架构	MOLAP 架构	HOLAP 架构	HOLAP 架构
查询性能	亚秒级查询延迟	亚秒级查询延迟	亚秒级查询延迟
并行计算引擎	HBase Coprocessor	HBase Coprocessor	Spark
存储引擎	HBase	HBase	KyStorage*
查询下压	○	●	●
明细查询	○	○	●
图形化设计器	○	●	●
智能辅助建模	○	●	●
LDAP/ Kerberos	●	●	●
单元格级别访问控制	○	●	●
内置敏捷 BI 工具	○	●	●
KyBot 诊断优化服务	○	●	●
技术支持	社区	24 X 7 SLA	24 X 7 SLA

* KyStorage 是 Kyligence 基于 HDFS 全新研发的拥有自主知识产权的列式存储引擎



开源承诺

Kyligence是由Apache Kylin的核心研发团队建立。KAP完全基于Apache Kylin核心代码，和开源版无缝兼容，并与开源社区协调一致发展，支持Apache Kylin的持续改进，多位Apache Kylin社区的committers均为KAP核心开发人员。Kyligence承诺持续高质量投入Apache Kylin的研发和社区发展，员工将持续为开源社区作出贡献。团队拥有Apache Kylin多位PMC（项目管理委员会成员）及核心committer。

Apache Kylin业界认可

致力于引领IT决策者走在科技前沿的国际科技媒体品牌InfoWorld，评选出的年度科技奖项Bossie Awards，根据软件对开源界的贡献，以及在业界的影响力评判获奖对象，被认为是当今IT界质量最高，最具创新能力的项目。Apache Kylin在2015、2016连续两年蝉联最佳大数据工具奖，与Google Tensor-Flow，Apache Spark、Apache Kafka，Apache Impala等荣登榜单，这是第一个来自中国的获奖项目。



“与Apache kylin团队一起合作使kylin通过孵化器成为顶级项目对我而言非常激动人心，Kylin在技术方面当然是振奋人心的，但同样令人兴奋的是Kylin代表亚洲国家，特别是中国，在开源社区中越来越高的参与度。”

—— Ted Dunning

Apache孵化器项目副总裁，MapR首席应用架构师

不积跬步
无以智千里



上海跬智信息技术有限公司

上海浦东新区亮秀路 112 号 Y1 座 405 室

info@kyligence.io

+ 86 - 021 - 61060928

<http://kyligence.io>

